

Weekly Report

基于排名学习的地图匹配方法

摘要: 从全球定位系统(Global Positioning System, GPS)获取的地理位置往往和实际的地理位置有一定的偏差。这制约了基于地理数据的数据挖掘、可视化、智能交通系统等方向的研究与应用。地图匹配是解决这一问题的方法之一,它将从GPS获取的地理位置匹配到对应的道路中。之前的地图匹配方法从几何、拓扑、概率等方面分析数据特点,建立数学模型来完成地图匹配任务。这些模型只能较好地反映研究时所用数据的特点而不能很好地反映具体应用场景的数据特点。这降低了这些方法的泛化性。为解决这一问题,本文提出了一种数据驱动的基于排名学习的地图匹配方法。该方法可直接从数据中学习地图匹配规则而不必建立具体的数学模型。据我们所知,这是世界上首次提出的从数据中直接学习地图匹配规则的方法。实验证明,我们的方法在XXX方面的有XXX效果,并在运算速度上相对部分算法有显著优势。

关键词: 地图匹配; 排名学习; 神经网络; 可视化; 地理信息系统

A Map Matching Method Based on Learning to Rank

Abstract: Omitted.

Key words: Omitted;

全球定位系统(Global Positioning System, GPS)是利用多颗定位卫星发射的定位信号,在全球范围内实现实时定位、导航的系统。近年来,随着电子技术的飞速发展,越来越多的设备装有GPS模块,人们也更容易使用GPS数据实现诸如导航、重要目标定位等基于地理位置的应用。另一方面,手机、传感器等设备的发展也使得更容易获取海量的GPS数据。海量的GPS数据为研究交通拥堵,行程预测,道路关联性分析等交通问题提供了新的研究途径^[2]。然而,由于设备性能及安全方面的原因,实际中往往无法获得准确、高频的GPS数据。卫星时间等参数的测量误差导致GPS获取的位置数据往往和真实位置有一至数十米的偏移;由于传输过程中数据的丢失,设备的流量及功耗限制,GPS数据的采样频率通常是不稳定且存在大量低频数据的^[9]。采样频率过低导致由连续两个GPS数据确定的位置之间的间隔过大,使得人们无法有效确定它们之间的轨迹。这类轨迹的不确定性在一些与GPS数据相关的研究中又称为采样误差^[1]。

未经处理的GPS数据的测量误差及采样误差限制了很多基于地理位置数据的研究和应用。比如,由于数据的测量误差和采样误差,无法确定每辆车在某时刻所在的具体道路,直接将轨迹可视化易造成轨迹与道路的重叠、混淆。地图匹配(Map Matching)则是解决该问题的数据预处理过程之一。地图匹配假设车辆或人行走在某条道路上而非道路外,将GPS数据确定的地理位置修正到其最可能在的道路上。地图匹配并非简单的问题。一方面,由于存在平行道路、隧道、Y型道路等复杂的城市地形^[3],往往难以确定与GPS数据对应的道路。另一方面,由于数据的规模和应用场景的多样性,实践中对算法的实时性、运算速度,对数据的存储、传输方式提出了不同的要求。

目前国内外已经有许多工作提出了不同的地图匹配方法。这些方法,就我们所知,绝大部分都是从几何、拓扑、概率方面分析数据的特点,建立相应的数学模型来完成地图匹配。因此部分文献^[3, 10-11]根据地图匹配算法所用到的技术的特点,将现有的地图匹配算法分为几何法

(geometry-based), 拓扑法(topology-based), 概率法(probability-based)和其它高级技术法。几何法从几何出发, 利用诸如点到道路的垂直距离等几何特征实现地图匹配。拓扑法利用道路网络的连通性等拓扑特征实现地图匹配。概率法则利用历史数据建立概率模型来完成地图匹配。高级技术法则从前三种方法出发, 引入启发式算法、机器学习模型等高级技术来完成地图匹配。典型的基于高级技术的地图匹配方法有基于隐马尔科夫模型(Hidden Markov Model)的方法和基于模糊理论的方法。基于高级技术的地图匹配方法也是目前学界研究最广泛的方法。抛开所用的具体高级技术不谈, 这些方法均是分析数据的几何、拓扑、概率的特点后, 建立相应的计算公式来实现的地图匹配方法。这带来了以下问题: (1) 需对 GPS 数据施加诸如采样频率等方面的限制; (2) 算法仅适用于部分目标(行人或车辆)和部分区域; (3) 参数选择问题。同时这制约了地图匹配更进一步的研究。

我们认为如果有种方法能够直接从数据中学习地图匹配规则(而非依靠人类知识)则能够解决依靠数学模型带来的局限性。然而目前大部分机器学习方法聚焦于解决分类、回归、聚类、降维等问题, 缺少一种朴素的方法使用机器学习模型实现数据驱动的地图匹配方法。我们受信息检索方面的启发, 提出了一种数据驱动的基于深度神经网络排名学习的地图匹配方法。该方法直接从数据中学习地图匹配规则而不建立具体明确的计算公式。该方法从数据中学习一个评分函数, 对 GPS 数据所有可能的道路进行评分, 选择评分最高的道路作为最终匹配的结果。据我们所知, 该方法是世界首次提出的直接从数据中学习匹配规则的地图匹配方法, 同时也是世界首次用深度学习来解决地图匹配问题。实验证明, 我们的方法取得了和目前已有的地图匹配算法相当的准确度, 并在运行速度上有显著优势。我们的方法可以通过选用更好的网络结构、优化目标来获得更高的准确度, 这部分的工作我们放在未来完成。

(我们的贡献: 设计了一个网络, 能用来完成评分函数。评分函数的训练则采用 rankNET loss。用来完成地图匹配算法)

1 相关工作-

(介绍地图匹配领域与排名学习(learning to rank)领域的研究进展。)

本节将简要介绍目前地图匹配和排名学习的相关研究。

1.1 地图匹配

(地图匹配研究进展。Data-driven 的地图匹配方法)

1.2 排名学习

(排名学习介绍。Point-wise, pair-wise, list-wise)

2 基于排名学习的地图匹配方法

(本节介绍我们提出的方法。)

2.1 数据及问题定义

(我们要解决的问题的 GPS 轨迹、路网、匹配结果、候选道路集、地图匹配等概念形式化的定义。)

2.2 网络结构

(输入、输出数据格式及网络的结构。)

2.3 损失函数

(训练所用损失函数及训练过程。)

3 实验结果及分析

(本节介绍验证我们方法有效性的实验。)

3.1 数据说明

(说明我们实验用到的数据。)

3.2 训练过程及评判标准

(说明我们的训练过程和评判标准。)

3.3 结果分析

(展示最终结果， 并做相应分析。)

4 结论

(总结全文。)

References :

[1] Hall T, Beecham S, Bowes D, Gray D, Counsell S. A systematic literature review on fault prediction performance in software engineering. IEEE Trans. on Software Engineering, 2012,38(6):1276–1304.

[2] Hunter, Timothy, Pieter Abbeel, and Alexandre Bayen. "The path inference filter: model-based low-latency map matching of probe vehicle data." IEEE Transactions on Intelligent Transportation Systems 15.2 (2014): 507-529.

[3] Zheng, Yu. "Trajectory data mining: an overview." ACM Transactions on Intelligent Systems and Technology (TIST) 6.3 (2015): 29.

[4] Lou, Yin, et al. "Map-matching for low-sampling-rate GPS trajectories." Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems. ACM, 2009.

[5] Newson, Paul, and John Krumm. "Hidden Markov map matching through noise and sparseness." Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems. ACM, 2009.

[6] Yuan, Jing, et al. "An interactive-voting based map matching algorithm." Proceedings of the 2010 Eleventh International Conference on Mobile Data Management. IEEE Computer Society, 2010.

[7] Hu, Gang, et al. "If-matching: Towards accurate map-matching with information fusion." IEEE Transactions on Knowledge and Data Engineering 29.1 (2017): 114-127.

[8] Rappos, Efstratios, Stephan Robert, and Philippe Cudré-Mauroux. "A force-directed approach for offline GPS trajectory map matching." Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM, 2018.

[9] Yuan J , Zheng Y , Zhang C , et al. An Interactive-Voting Based Map Matching Algorithm[C]// Eleventh International Conference on Mobile Data Management, MDM 2010, Kanas City, Missouri, USA, 23-26 May 2010. IEEE, 2010.

[10] Quddus M A , Ochieng W Y , Noland R B . Current map-matching algorithms for transport applications: State-of-the art and future research directions[J]. Transportation Research Part C Emerging Technologies, 2007, 15(5):312-328.

[11] Ta N , Wang J , Li G . Map Matching Algorithms: An Experimental Evaluation[C]// Asia-pacific Web. Springer, Cham, 2018.

工作时间

星期	任务	Duration
周一至周日	调试代码，看文论，构思论文，写论文	9:00 - 12:00 和 14:00 - 22:30，共 11 小时

Work Time: More than 50 hours